Security for Digital Health Care: What's There? What Works? What's Needed?

S. Demurjian[1], A. De la Rosa Algarín[1], J. Bi[1], S. Berhe[1], T. Agresta[2], X. Wang[2], and M. Blechner[2]

[1] Department of Computer Science & Engineering
University of Connecticut, 371 Fairfield Rd., Unit 4155
Storrs, CT, 06269-4155 USA
{steve, ada, jinbo, solomon.berhe}@engr.uconn.edu
[2] Departments of Family Medicine and Pathology
University of Connecticut Health Center, 263 Farmington Avenue
Farmington, CT, 06030 USA
agresta@nso1.uchc.edu, {xiaowang, mblechner}@uchc.edu

Abstract. Today's systems use a myriad of data formats (e.g., XML, RDF, JSON, SOAP, etc.) for data exchange and sharing. This is particularly true for health care, where patient information is stored in different locations via electronic media and/or hard-copy formats. Such information is of interest to: providers (MDs, RNs, therapists, etc.) to support patient care; researchers interested in secondary usage of data for data mining on medical conditions; public health researchers looking at population data across a region; insurers tracking patients with similar conditions and analyze delivery of care; and, patients who access health data in emergent situations. All potential users need electronic access to health information technology (HIT) systems such as: electronic health records; personal health records (PHRs), whose information availability is controlled by patients; patient portals to request appointments, refills, and referrals; and, ancillary systems such as imaging, laboratory, pharmacy, etc. Controlling access to information from multiple HIT systems require granularity levels of privileges ranging from one patient to a cohort to an entire population. For secure access to patient data, security policies and enforcement mechanisms will need to consider: a patient consent for a provider to access data in a PHR; a patient's data stored in multiple locations for direct patient care; a group of patient data collected from multiple HIT systems to study specific disease onset, progression, treatments, and their outcomes; and, a large de-identified data set for disease surveillance and public health research. To accomplish these types of secure access, we need to know what security alternatives are available, what resources need to be protected, what approaches currently work in the present state-of-the art, and most importantly, what's potentially needed for the future of security for digital health care (DHC).

Keywords: Digital health care, health information technology, security, health information exchange.

1 Introduction

As we look back over twenty years ago, we recall that two articles related to health care security were published in the IFIP WG11.3 3rd conference. In [1], privacy and confidentiality in medical information systems was explored, advocating a role-based approach, and detailing the state-of-the-art in available systems. In [2], a case study of mental health delivery from information and semantic perspectives was presented, providing scenarios of usage of information by physicians, nurses, etc., and promoting a role-based approach as the most appropriate solution. What is surprising is what has stayed the same and what has changed over 22+ years in the health care in terms of tracking patient care (via paper or electronic form) and facilitating secure information exchange as a patient transitions between care settings. For instance, in 1990, would anyone have predicted the Health Insurance Portability and Accountability Act[3] of 1996 (HIPAA) Privacy and Security Rules for protected health information (PHI), that at the time was based more on paper than electronic health records (EHRs), the Genetic Information Nondiscrimination Act (GINA)[4] of 2008 that protects a patient's genetic information against discrimination in health insurance and employment, or the Ethical, Legal and Social Implications (ELSI) research program to mange genomic data for personalized medicine? There have also been dramatic changes in patient care, including: EHRs in some MD offices ("implementation rates reached 68% in family practices in 2011"[5] while "just 27% of physicians used EHRs with multi-functional capabilities"[6]); personal health records (PHRs) for patients to store their own health information, download medications from a pharmacy, and potentially share data with providers; a Patient Centered Medical Home (PCMH)[7] where one provider coordinates care for patients with chronic diseases; an accountable care organization (ACOs)[8] to coordinate providers regarding Medicare patients with chronic conditions; or clinical research data warehouses (CRDW) which are de-identified repositories that support clinical and population research, etc., via the collection of data from multiple health information technology (HIT) systems.

The harsh reality in health care and HIT adoption is the limited capabilities of health information exchange (HIE) among all of these various data sources, the high number of providers that are predominately paper based with limited or no access to EHRs or other HIT systems, and most importantly, the fact that security is often an afterthought in this process, supported for individual systems for specific providers, but overlooked when one attempts to bring together patient

[3] http://www.hhs.gov/ocr/privacy/
[4] http://www.genome.gov/24519851
[5] http://www.aafp.org/online/en/home/publications/news/news-now/practice-professional-issues/20130201ehradoptrates.html
[6] http://www.informationweek.com/healthcare/electronic-medical-records/ehr-adoption-us-remains-the-slow-poke/240142152
[7] http://www.pcmh.ahrq.gov/portal/server.pt/community/pcmh__home /1483
[8] http://www.innovations.cms.gov/initiatives/ACO/index.html

data from multiple electronic sources in support of PCMH, ACOs, CRDWs. In PCMH, the effective care of a diabetes patient with high blood pressure may involve the family practitioner (who sees the patient regularly), an endocrinologist (if diabetes is complex in its manifestation), a cardiologist (if there are heart issues), and a nutritionist (for managing diet or dealing with obesity). These four providers may have different EHRs (or none) and an inability to share data (patient history, lab test results, etc.) to facilitate the required care. The access needs to be integrated (electronic sources), secure (individual sources and across the integrated sources), and collaborative (individuals can view/update same patient record simultaneously). Our main objective in this paper is to enumerate prevelant issues for secure, integrated, and collaborative health care, requiring us to provide a roadmap for digital health care (DHC) in the not so distant future that includes, but is not limited to answering questions such as: what patient information is available for each source, how is the local security for that source managed, what needs to be protected from each source, is there a global security policy across the integrated sources, what security methods in the present state-of-the-art are appropriate to employ.

The remainder of this paper has five sections. Section 2 presents a high-level view of security needs for DHC focusing on laws, standards, and emerging models of care spanning clinical, genomic, and phenotypic information. Section 3 provides a scenario on the actual experiences of one co-author in navigating the health care system with HIT in use at some level by most providers, but with paper-based records still exchanged, snail mail and fax. Section 4 details a proposed security framework that considers all of the constituent HIT systems, standards, and applications and their interactions. Using this as a basis, Section 5 proposes a core set of recommendations organized by area that that must be supported for security for DHC. Finally, Section 6 concludes the paper.

2 High Level View: Security for DHC

Security for DHC goes well beyond the needs of compliance of HIPAA, which provides a set of security guidelines in the usage, transmission, and sharing of PHI. In addition, there is a need to: protect personally identifiable information (PII), including names, addresses, accounts, credit card numbers, etc.; encrypt PHI and PII data and its secure transmission (e.g., SSL); extensive usage of standards for storage and exchange (Health Level Seven (HL7) clinical document architecture (CDA)[9] and the Continuity of Care Record[10] (CCR) for administrative, patient demographics, and clinical data); leveraging a wide range of health care standards (e.g., CDA, CCR, LOINC[11], SNOMED[12], UMLS[13]),

[9] http://www.hl7.org/
[10] http://www.astm.org/Standards/E2369.htm
[11] http://loinc.org/
[12] http://www.ihtsdo.org/snomed-ct/
[13] http://www.nlm.nih.gov/research/umls/

and dealing with data interoperability issues for HIT systems that use a wide range of data formats (e.g., XML[14], RDF[15], JSON[16], etc.). Instead, to attain security for DHC, we will need all of these underlying technologies and standards coupled with a strong understanding of the way that health care data is utilized by patients, providers, researchers, etc. We must also include the emerging need to manage genomic data for personalized medicine and its potential future integration and/or consolidation with EHRs via the Ethical, Legal and Social Implications (ELSI) that is tied to the Genetic Information Nondiscrimination Act (GINA) of 2008. GINA protects a patient's genetic information against discrimination in health insurance and employment, including: genetic test of patient, his/her family members, fetus of individual or family member, family medical history, and request/receipt of genetic services that may include research trials[17]. HIPAA's rule insures that PHI is securely maintained with patients retaining rights to their information stored in a PHR (patient controlled) or to access information from a provider's record (EHR or hard copy), while still allowing entities to disclose the information under certain situations. HIPAA's rule defines the "series of administrative, physical, and technical safeguards for covered entities to use to assure the confidentiality, integrity, and availability of electronic protected health information". For ELSI, protection of information must be reconciled across HIPAA and GINA to securely deliver the appropriate combination of clinical, genomic, and phenotypic information to researchers, clinical providers, support personnel, insurers, and patients.

Security for DHC transcends just protecting the information, and must strongly consider the usability of the information by a wide variety of stakeholders using a broad range of HIT systems to effectively and securely leverage different types of patient data. The HIT systems available include: EHRs, electronic repositories of patient medical records that may exist in provider offices, clinics, and hospitals; PHRs such as MS HealthVault and webMD that allow patients to manage their own health care data and decide which provider(s) to allow access to that information; patient portals that allow patients to electronically request appointments, prescription refill requests, arranging a referral to another provider or specialist; personalized medicine health Portals (PMHPs) such as Genomas [18] which allows providers to view their own patients' genetic data against their medical record (EHR) in order to bridge the gap between providers and medical researchers; ancillary systems for laboratory results (e.g. blood work), evaluating X-rays, MRIs, CT Scans, etc., electronically so that reports and images can be delivered to providers, pharmacy systems for tracking medication and interactions, etc.; Patient applications for access to education information and management of chronic diseases, medications, and interactions with providers; and, a clinical research data warehouse (CRDW) that contains de-identified clin-

[14] http://www.w3.org/XML/
[15] http://www.w3.org/RDF/
[16] http://www.json.org/
[17] http://www.genome.gov/Pages/PolicyEthics/GeneticDiscrimination/GINAInfoDoc.pdf
[18] http://www.genomas.net/

ical data loaded from medical records for patients to have their data used for medical research, or for public health researchers to do population studies.

All of these systems target a wide range patient care and research initiatives. First, a Patient Centered Medical Home (PCMH) can manage chronic conditions and optimize care by interacting stakeholders (e.g., family practitioner, endocrinologist, cardiologist, and nutritionist example in Section 1); in this situation, there may be a need for the lead provider to access information in other EHRs, PHRs, etc., in a timely manner in order to coordinate effective care. Second, accountable care organizations (ACOs) bring together larger groups of providers, clinics, hospitals, and private insurers in an effort to give coordinated care to a panel of medicare patients in order to attempt to reduce or eliminate duplicate test and procedures for patients that visit multiple providers and have chronic conditions. Third, secondary use (SU)[19] of clinical data allows providers and researchers to analyze specific diseases and their treatments across a large patient base for example via a clinical research data warehouse (CRDW), seeking events such as adverse drug reactions, infection monitoring, disease monitoring in a larger population (the flu epidemic in the United States in 2013), etc.. Fourth, meaningful use (MU)[20] which is focusing on the adoption and use of HIT within organizations that may lead to improvements in the reporting of care by offering providers incentives to acquire and deploy technology. Fifth, personalized medicine (PM)[21], which is targeting the treatment of an individual based on their unique medical profiles that might include specific types of diseases and focus on the use of a patient's genomic information.

In support of these aforementioned initiatives, health information exchange (HIE) is vital to insure that the correct data is available at the appropriate time in a usable fashion by a specific stakeholder. Note that what is shared by HIE is most often determined by the institute that owns the data; it doesn't mean all of the data is shared and the data to be shared is often off-loaded into another server intended for that purpose so that there is no impact on the real-time usage of an EHR. For example, HIE allows sharing so that in emergent situations providers can retrieve data on a patient from a system they are not authorized via a techniques such as dynamic certification. Alternatively, HIE can be used to construct a CRDW via an electronic, transfer, load (ETL) process from an EHR database, which can provide, for example, workflows and ontologies for managing tissue data including controls for patient consent relating to tissues and boundaries on experimental uses. HIE and other means of extracting clinical and claims (and other) data can also be utilized to support analysis for SU, ACO, and MU, providing de-identified data to clinical researchers so that best practices can be evaluated across a wide range of clinical settings. Our intent in this paper is to consider all of the above factors in order to propose an architecture that ideally achieves security across this entire spectrum of standards,

[19] http://www.ncvhs.hhs.gov/050726p5.pdf
[20] http://www.hhs.gov/news/press/2010pres/07/20100713a.html
[21] http://www.personalizedmedicinecoalition.org/

regulations, HIT systems and their usage by stakeholders, coupled with HIE and supporting a wide range of data analyses.

3 Low-Level View: Intelligent Security in DHC

To better understand health care and the impact of HIT on patient care, this section along with Figure 1 provides a realistic case study of one co-author navigating through a complex process. Consider that a 54 year old man falls while working in the yard, and breaks his hip; an ambulance (Step 1 in Figure 1) takes him to the ER of a small regional hospital (Step 2) where a history is taken using an EHR at the hospital, X-rays are ordered, and a hip fracture is found. After speaking with the ER MD, and talking to a physician colleague, the patient decides to transfer by ambulance (Step 3) to a metropolitan area hospital, and his records are sent in hard copy. Upon arriving at the ER of that hospital, another patient history is taken for that hospital's EHR to capture the same information. The same story is told to ER MD, orthopedic resident, etc., and at 2am in the morning he signs a consent form for either a partial or full hip replacement. At 7am the transport team arrives to take him to the OR; at that point, the orthopedic surgeon has another option, to repair the hip with a plate and screws, and the patient, after consulting with his physician colleague, has to re-initial the hardcopy consent form. Surgery is successful, and after three more days in the hospital, the patient is discharged to a rehab facility, (Step 4) again with a hard copy of his records. The rehab center is mostly paper-based; they have an electronic system, but, for example, the medication list is hard copy as the nurse dispenses meds to patients. After 5 weeks, the patient is discharged (Step 6) to his home, and the Visiting Nurse Association in his area is assigned to monitor his care via a nurse and in home physically therapy.

During the time at rehab and at home, the patient visits the surgeon (Step 5 and Step 7) in order for x-rays to assess the healing, and also meets with his internist (Step 8) for follow up care. The internist has an EMR and can download all tests done at an external lab facility, but records at the hospital will have to be faxed and then scanned and put into the EHR as images (unsearchable). Ten weeks after the fracture, the patient is given his release from the orthopedic surgeon (Step 7) with weight bearing, and asks for that care to be managed by a local orthopedist (Step 10). The patient requests medical records be sent to the local orthopedist, but 2 weeks later at the appointment, no records have arrived, appointment, so new and old X-Rays can't be compared. Due to the unusualness of a hip fracture of a 54 year old, the patient is referred to a rheumatologist (Step 9), and brings a hard copy of some of his medical records to the appointment; blood work and a bone scan that determines that the patient has osteoporosis. The rheumatologist's office has no EHR, but can access systems at an imaging facility and testing laboratory; the rheumatologist's also makes a medical record request from hospital. Consider that even with the advance of technology and its availability, fax and snail mail are still playing a dominant role in the way that

we transfer healthcare data. How can a rheumatologist without an EHR get all of the information needed from multiple sources in a timely fashion so as not to delay treatment? Clearly, even if we can deal with security for DHC, there will still be a huge hole in the overall security of patient data with information in so many different and incompatible locations and continued dependence on paper.

Further, suppose that a clinical researcher was interested in conducting a study of males 50-60 who have hip fractures and osteoporosis, and what they may have in common (e.g., low vitamin D, low testosterone, low calcium, etc.). The dramatic push to digitize clinical data via EHRs has led to an unprecedented opportunity for clinical and public health studies [3,4,5,6]. This growth is being fueled by recent federal legislation that provides generous financial incentives to institutions demonstrating aggressive application and meaningful use of comprehensive EHRs [3]. Efforts are already underway to link these EHRs across institutions, and standardize the definition of phenotypes for large scale studies of disease onset (as well as rate of progression) and treatment outcome, specifically within the context of routine clinical care [7,8,9]. The longitudinal nature of the data contained within EHRs makes them ideal for quantifying outcomes from the utilization of prescription medications (both efficacy and toxicity). More recently, huge efforts have been initiated to link new and existing EHR databases to accelerate research in personalized medicine [7]. This is a herculean task in most of the clinical environment without an established informatics infratructure, since to find enough of a patient cohort, the researcher would need to query multiple hospitals, surgeons, laboratories, internists, and rheumatologists. At the present time, HIE has not advanced to a stage to support such queries in any reasonable time frame. In such a scenario, how can the security issues that span multiple HIT systems each with their own security control (with local HIPAA compliance) be brought together to securely obtain this data (with a more global HIPAA compliance) into a de-identified CRDW to facilitate the research? How is data securely gathered into this CRDW from paper sources? How is institutional review board approval obtained when the patients may be from multiple institutions? How is HIPAA compliance of hard-paper copies at physician offices that are transferred via fax and/or snail mail protected until they are entered into the CRDW repository?

Security for DHC must anticipate a future where the medical community has caught up with the use of HIT, and must consider EHR vendors that do not wish to allow their information to be easily shared, as do hospitals, since they deem sharing of data to cause the potential for loss of patients to other hospitals. The EHRs for the regional and metropolitan hospitals do not share data, and may not share data with local providers (e.g., internists, rheumatologists, local orthopedist, etc.). Do we define a solution with the expectation that we are planning for a futuristic scenario where secure sharing and exchange is the norm and HIT is in almost all providers? Is this even realistic in today's medical system in the US or even within the next 5 years? 10 years?

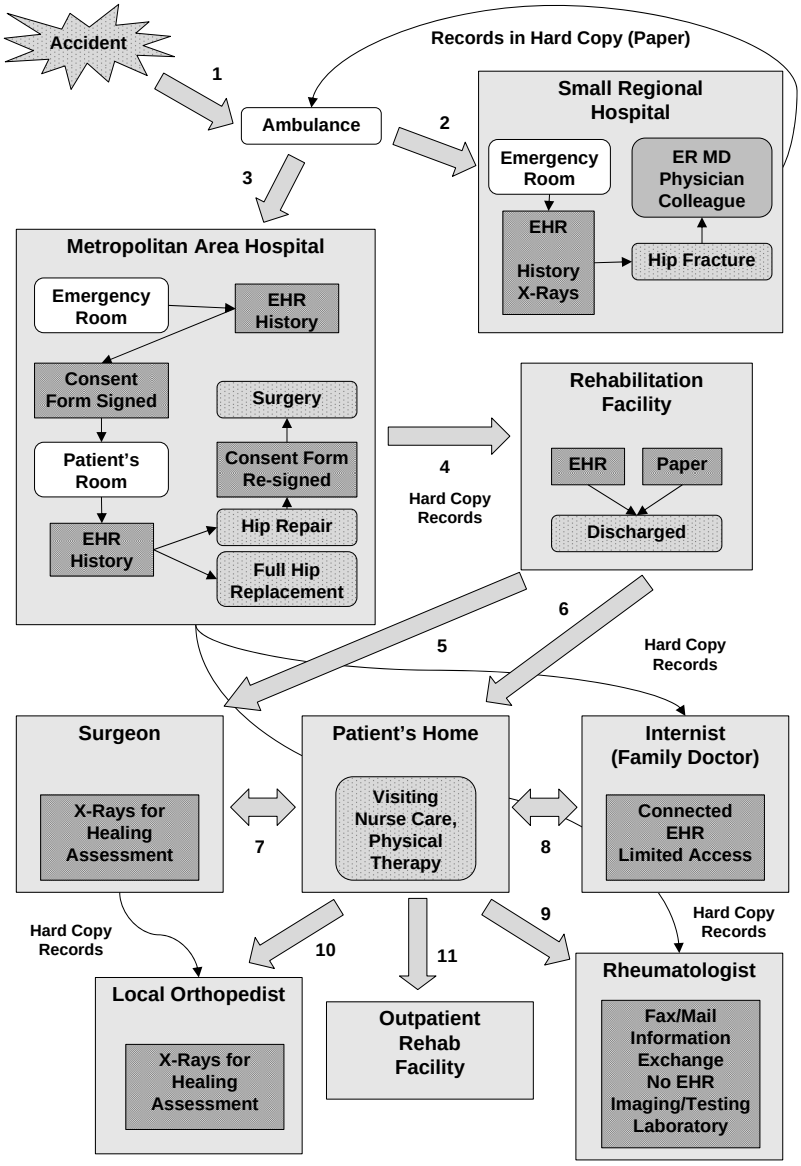

Fig. 1. Illustrating a Sample Health Care Process

4 Proposed Architecture for Secure DHC

For successful HIE, the security of constituent systems must be integrated and support the application's need. What happens when security privileges of individual systems are in conflict with one another? How do we reconcile these local security policies? Is it possible to define a global encompassing security program providing a level of guarantee to the local security policies from an enforcement perspective? As today's health care applications continue to become more complex and wide-spread, interacting with many other systems (or applications) using varied technologies, there is a need for some degree of assurance that security for the application (global) is consistent with the sum of the parts (local security) of the constituent systems.

To place our work into its perspective, Figure 2 shows the HIT systems, XML medical standards, and end-user applications, to provide an infrastructure for PCMH, ACO, SU, MU, and PM. We examine the infrastructure from three viewpoints: the reconciliation of security (local and global) to insure that the required clinical data reaches the providers involved in PCMH and PM where data mining and knowledge discovery techniques can be used; the availability of de-identified patient data for providers and clinical researchers via privacy-preserving data publishing and sharing in support of ACO, SU, and MU that can be used for data analysis, mining, and clinical decision support to learn what works and what doesn't in terms of treatments of various illnesses and diseases; and, the facilitation of the first two perspectives via the use of XML and associated standards for patient and clinical data (CDA, CCR, etc.), and ontologies that augment this data with relevant tags that add meaning (SNOMED, LOINC, NDF-RT, etc.). The lower left of Figure 2 contains EHR systems (VistA, openEHR, PatientOS, GE Centricity, FreeMED) that all share an ability to export patient data, in XML formats or standards (generic XML, CCR, and HL7's CDA), via the use of a proxy server in which data from the EHR has been offloaded. Emerging platforms (Open mHealth, to promote mobile health via an open architecture, and the Harvard SMART platform for substitutable medical applications that promote reuse) and PHRs (Microsoft HealthVault and EPIC LUCY) are presented in the upper-left of Figure 2. Open mHealth uses JSON to model patient data, while SMART uses RDF/SMART and JSON-LD. These choice of data formats must be converted (XML-C) before they can be secured.

The bottom right contains examples of medical applications that must be securely managed, PHA and SMARTSync. Personal Health Assistant (PHA) is an in-house developed mobile (not publicly available), test-bed application for health information management that allows: patients to view and update their personal health record stored in their HealthVault account and authorize medical providers to access certain portion of protected health information; and, for providers to obtain the permitted information from their respective patients that they have been authorized to view. The patient version of PHA allows users to perform a set of actions regarding their health information. Users can view

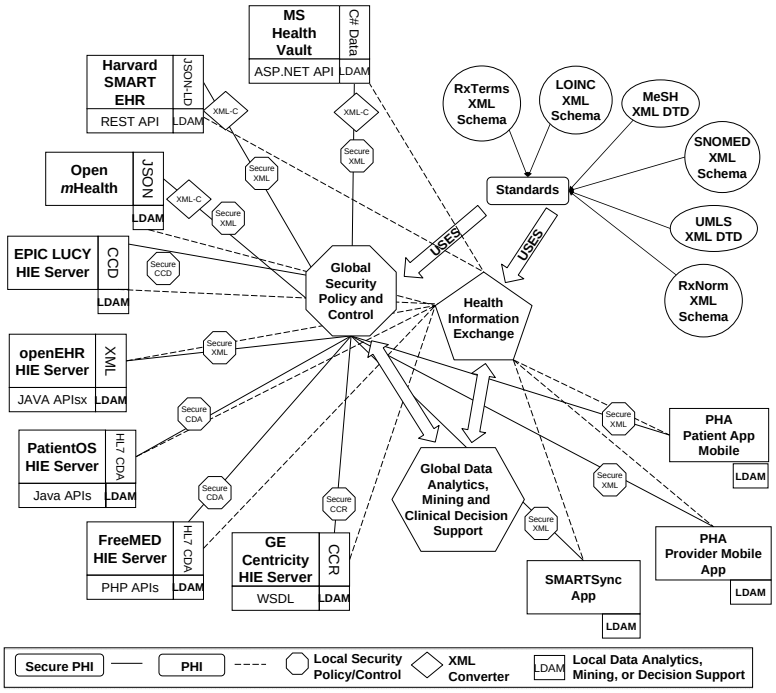

Fig. 2. Proposed HIE Architecture with Information Security and Analytics

and edit their medication list, allergies, observations of daily living, and set security policies for read/write permissions on their medical providers by role as reported in our prior work [10]. Security settings can be set at a fine granular level, and each provider gets view/update authorizations to the different information components available in PHA. The provider version of PHA allows the users (health professionals) to view and edit the medical information of their patients as long as there are permitted to do so as dictated by the security set by the user (patient). SMARTSync is an in-house developed (not publicly available), web-based test-bed medication reconciliation application used to create and preserve a patient's medication list through transfers among locations of care, preventing immediate interactions, and avoiding dosage errors in situations where brand and generic drugs are received or multi-component drugs are used. Significant risks include: overmedication when a provider prescribes a new medication (or one from the same class) or when an interacting medication is prescribed; adverse interactions, the result of conflicts between medications, which can change effect strength or serum concentration; and adverse reactions, allergic/other effects, experienced by patients which can result in a patient being wrongly labeled as allergic to a medication, unnecessarily excluding it as a treatment option in the future. To accomplish this, we gathered data form HealthVault and SMART Reference EHR [11].

The upper right contains the various standards for medical information, such as RxTerms and RxNorm (services to augment medication information), medical codes (SNOMED), medical nomenclature (UMLS and MeSH), and laboratory codes (LOINC) that are used by HIT systems and applications. In addition, local data analytics, data mining, or decision support (blue square component of systems) can be found in institutional HIT systems, patient PHR solutions or end-point applications (e.g., SMARTSync, PHA). This data analytics component exists in a globalized manner, where researchers external to all of the local components will need access to information found a distributed HIT systems.

The three main aspects that allow all the interaction to occur across Figure 2 are presented by HIE (pentagon), that uses dotted lines to indicate the necessity to share data among HIT end-points; the Global Security Policy and Control (octagon), that provides a centralized representation of which interactions can locally occur; and, Global Data Analytics, Mining and Clinical Decision Support (hexagon), which in conjunction to the local counterparts (LDAM in each respective system), provide the communication between researchers who seek data to discover hidden knowledge (on a global perspective), or from the local system's point of view (perspective), provide some sort of intelligent support mechanism, like PHA does with medication interaction checking. The end result and major challenge presented in Figure 2 is the recognition of a greater need for a comprehensive approach to security at global and local levels operating an environment that is driven to share data through HIE. This comprehensive approach needs to take into consideration the distributed nature of repositories, the fragmentation of patient data across these systems, the discrepancies on sharing and security policies set in each component, as well as the potential usage of parties that do not own the data and are merely borrowing it in a predetermined set of constraints (e.g., time, amount, demographic, etc.).

Data mining techniques have long become useful tools in solving security related problems [12], permitting the extraction of information or knowledge from collected data or observed examples using statistical algorithms. Data mining methods have applications in intrusion detection, insider analysis and many others [12,13]. Particularly, for the proposed security infrastructure (Figure 2), the reconciliation of security policies in both global and local levels may benefit from mining large-scale data transition-recording files or security monitoring log files. The knowledge patterns detected from the mining steps may bring insights into a revision of the existing policies and a better reconciliation. For instance, cloud computing may become a necessary resource for HIE, and insider problem has been cited as the most serious security problem and the most difficult problems to deal [14]. In most organizations, insider problems were limited to authorized internal employees. However, in cloud computing context, insiders may be expanded from organization internal employees and contractors to cloud internal employees and contractors, cloud customers, and cloud third party suppliers. This expansion increases the exposed threats on healthcare organizations sensitive data, such as PHI that is being transitioned between and shared among different organizations. Data mining approaches such as cluster analysis, novelty

detection and association rule mining can be used to examined the data-access log files and detect abnormal patterns in various transactions.

The practice in knowledge discovery from large compiled data also imposes great challenges to security, especially during the process of sharing EHR data. The national and state healthcare agencies routinely publish EHRs for secondary data analysis that aims to expand knowledge about disease and treatments and enhance healthcare experience for individuals. The access and aggregation of EHRs poses significant concerns about patient privacy and confidentiality. According to HIPAA, de-identified healthcare information may be used and disclosed for secondary analysis. "De-identified" is defined that the personal identifiers in a record have been extracted and it is difficult to re-link the data to the people mentioned in the original records. "Anonymized" means that all of the links between a person and the person's record have been irreversibly broken so that it would be impossible to re-identify the person in the records. However, in large-scale secondary analysis of multiple data sources that involve race, ethnicity, gender, service date, diagnosis codes (ICD9), or procedure codes, by cross linking these data sets with other publicly available databases, data mining methods may be able to associate an individual with specific diagnoses. For example, it was demonstrated that an individual could be re-identified by linking certain attributes in a published data set with a voter registry (Cambridge, MA [15]).

5 Recommendations for Secure DHC

Our major assumption in this section is that a significant barrier to integrated patient care data access occurs when a stakeholder needs to access information from a HIT system s/he has not been previously authorized to use, in either a routine or emergent situation, and are not easily authenticated to access information from systems that they have not been previously authorized to. We recognize that in order to provide proper security, any recommendations must cover the storage, transmission and storage of PHI and PII data. In this section, we provide the different recommendations across a broad range of system techniques and mechanisms, as well as approaches that require a "person" in the loop in order to monitor and control secure information access.

Encryption: The distributed nature of data storage in healthcare makes it necessary to provide security at storage point, as well as in the point of transmission. An encryption framework must provide a robust level of security for stored information capable of integrating heterogeneous local solutions, in the respective data sources, in a global context. This encryption framework should be extensible to handle new types of data unique to health care (genomic, phenotypic). For secure online data transmission, existing technologies (e.g., HTTPS, SSL, etc.) should be leveraged in order to provide a proper level of protection. The HITECH Act achieves PHI portability and storage through encryption as applied to hard drives and (portable) systems such as laptops, jump drives, desktops, smart phones, tablets, cloud inter-system links, and user-system links [16].

Certificates: X.509 certificates and their ability to be extended via certificate attributes can allow, over time, a user to acquire multiple X.509 certificates (each to access a specific system) based on their activity being authorized to utilize different systems. The advantage of multiple certificates (one per work setting) is to minimize the impact for failure; with a single certificate and multiple attribute certificates (one for each work setting) failure may compromise multiple settings, while multiple certificates (one per work setting) should limit the impact of failure. Each work setting can have their own security infrastructure and algorithms to generate public-private key; the concept of multiple certificates each with multiple ACs attached is akin to a wallet with multiple cards issued from different sources [17]. Related efforts include: a framework for secure e-Health authentication using a multiple factor approach where physicians would provide multiple pieces (e.g., ACs) of information in emergent situations akin to our multiple certificate approach [18]; and, a framework for adaptive trust negotiation that establishes trust based on attributes other than identity [19].

DIRECT and Health Information Service Provider (HISP): DIRECT20 allows individuals, providers and organizations to share information with best practices that have trust and privacy considerations that are very consistent to the privacy emphasis of this proposal. HISP is used to describe the management of security and transport for directed exchange and an organizational model that performs HISP functions to allow interactions of HIPAA Covered Entities with the sender or receiver of directed exchange of PII, and must include all data collection, use, retention, and disclosure policies. In practice, sender and receiver take sole responsibility for encryption/decryption activities through the use of standardized encryption algorithms. In Figure 2, there would be HISPs for each of the data sources (EHRs and PHRs). HISP could use X.509 certificates as previously defined, where a certificate by role could be established for different roles. Attribute certificates can be associated with various characteristics such as for the data level (HIPPA, FERPA, DE-IDs), the situation (Urgent care, Primary Care, Inpatient Care), the type of data (patient, genomic, deidentified), etc. In an emergent situation where an MD needs access to another EHR, s/he could present her/his X.509 certificates and a process can be initiated by the to consult among the EHRs with two possible results: access is allowed to the MD based on submitted certificates (with some expiration) or not.

Access Control: Access control models provide the benefit of applying security at different levels of the information exchange scenario. Given the structure of Figure 1, role-based access control [20] could be used as a cornerstone, but needs extensions for health care. Extension parameters include patient, healthcare facility, task, temporal information, stakeholders [21,22]. Another extension would be the ability to extract local security policies and integrate them into a global one that is enforceable across the health care enterprise [23,24,10]. A third extension would be for delegation of authority to facilitate access in an HIE setting [21], which a provider often passes on his/her permissions (e.g., patients) to other providers. A fourth extension is the need for secure HIE across a wide range of data (for example, clinical, genomic, and phenotypic) which will involve

the co-consideration of HIPAA, GINA, and ELSI, and the exploration of RBAC and delegation for genomic and phenotypic data.

User-Based Security Mechanisms: There are many security nets in health care in terms of data access that happen after the misuse event has occurred. A clinician role in a hospital would have specific permissions, and an actual user should be further restricted to his/her patients. In practice, clinicians may be able to access more data than they are authorized and monitored in each single system against suspicious patient data retrieval [25], and this is done after the fact via an audit. However, as given in Figure 2, the detection of intruders or system misuse is going to be necessary and will require more sophisticated network monitoring tools against consolidated log files from all constituent HIT systems that are interoperating. One dominant approach for data access for HIE or CRDW is the use of an honest broker, an actual individual who is in charge of triggering the clinical (research) data request event to the corresponding HIT system and returning the results to the clinical (researcher) [26]. Large hospitals require a dedicated team of patient privacy security officers in charge of enforcing regular password updates, system updates, correct system configuration, hard drive encryption, and other security related tasks; clearly this is more complex in an environment as shown in Figure 2. Often, improving user access means that clinicians must be educated on privacy regulations, procedures, system usage and configuration, to avoid misconfigurations, such as using the same password for the private key, operating system login, and EHR system login that can merely achieved during dedicated seminars [27].

6 Conclusion

This paper has looked backwards to the first discussions of privacy and access control for medical settings [1,2], and more importantly forward to the wide array of emerging HIT systems, applications, and standards, intended to support HIE in order to allow varied stakeholders to securely access information in routine and emergent situations. As a result, security cannot be considered simply from these individual systems, but must take a approach that requires a more global security solution to protect the vast amount of data available for use by medical professionals and data analysis by researchers. Toward this objective, in this paper, Section 2 presented the changing landscape of medical care, standards, and technologies, that are difficult to support without HIE, and is further complicated by the present state of medical information exchange, as illustrated by a scenario of patient care recently experienced by one of the co-authors and detailed in Section 3. Based on this information, in Section 4 and Figure 2, we presented an architecture that positioned the HIT systems, standards, and applications in the context of HIE, global security enforcement, and global data mining and analysis. Using this as a basis, the proposed recommendation list in Section 5 is a significant starting point or roadmap for considering the security for digital health care that transcends individual systems and must consider

the diverse HIT systems, applications, standards, and their interactions, using and extending traditional security mechanisms (encryption, access control, auditing, etc.) and user based techniques with privacy officers to control access to information via honest brokers.

References

1. Biskup, J.: Protection of privacy and confidentiality in medical information systems: Problems and guidelines. North-Holland (1990)
2. Ting, TC: Application information security semantics: A case of mental health delivery (1990)
3. Shea, S. and Hripcsak, G.: Accelerating the use of electronic health records in physician practices. New England Journal of Medicine **362**(3) (2010) 192–195
4. Wang, X. and Chused, A. and Elhadad, N. and Friedman, C. and Markatou, M.: Automated knowledge acquisition from clinical narrative reports. In: AMIA Annual Symposium Proceedings. Volume 2008., American Medical Informatics Association (2008) 783
5. Wang, X. and Hripcsak, G. and Markatou, M. and Friedman, C.: Active computerized pharmacovigilance using natural language processing, statistics, and electronic health records: a feasibility study. Journal of the American Medical Informatics Association **16**(3) (2009) 328–337
6. Jha, A.K. and DesRoches, C.M. and Campbell, E.G. and Donelan, K. and Rao, S.R. and Ferris, T.G. and Shields, A. and Rosenbaum, S. and Blumenthal, D.: Use of electronic health records in US hospitals. New England Journal of Medicine **360**(16) (2009) 1628–1638
7. McCarty, C.A. and Wilke, R.A.: Biobanking and pharmacogenomics. Pharmacogenomics **11**(5) (2010) 637–641
8. Pace, W.D. and Cifuentes, M. and Valuck, R.J. and Staton, E.W. and Brandt, E.C. and West, D.R. and others: An electronic practice-based network for observational comparative effectiveness research. Annals of internal medicine **151**(5) (2009) 338
9. Ritchie, M.D. and Denny, J.C. and Crawford, D.C. and Ramirez, A.H. and Weiner, J.B. and Pulley, J.M. and Basford, M.A. and Brown-Gentry, K. and Balser, J.R. and Masys, D.R. and others: Robust replication of genotype-phenotype associations across multiple diseases in an electronic medical record. The American Journal of Human Genetics **86**(4) (2010) 560–572
10. De la Rosa Algarín, A. and Ziminski, T. B., and Demurjian, S. A., and Kuykendall, R., and Rivera Sánchez, Y.: Defining and Enforcing XACML Role-Based Security Policies within an XML Security Framework. In: Accepted, Proceedings of 9th International Conference on Web Information Systems and Technologies, INSTICC (2013)
11. Ziminski, T.B. and De la Rosa Algarín, A. and Saripalle, R. and Demurjian, S. A. and Jackson, E.: SMARTSync: Towards Patient-Driven Medication Reconciliation Using the SMART Framework. In: International Workshop on Biomedical and Health Informatics. (2012) 806–813
12. Lin, T.Y. and Hinke, T.H. and Marks, D.G. and Thuraisingham, B.: Security and data mining. Database Security **9** (1996) 391–399
13. Zhu, D. and Premkumar, G. and Zhang, X. and Chu, C.H.: Data Mining for Network Intrusion Detection: A Comparison of Alternative Methods*. Decision Sciences **32**(4) (2007) 635–660

14. Khorshed, M.T. and Ali, AS and Wasimi, S.A.: Monitoring insiders activities in cloud computing using rule based learning. In: Trust, Security and Privacy in Computing and Communications (TrustCom), 2011 IEEE 10th International Conference on, IEEE (2011) 757–764
15. Sweeney, L.: k-anonymity: A model for protecting privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems **10**(05) (2002) 557–570
16. Kwon, J. and Johnson, M.E.: Security practices and regulatory compliance in the healthcare industry. Journal of the American Medical Informatics Association (2012)
17. Mavridis, I. and Georgiadis, C. and Pangalos, G. and Khair, M.: Access control based on attribute certificates for medical intranet applications. Journal of Medical Internet Research **3**(1) (2001)
18. Boonyarattaphan, A. and Bai, Y. and Chung, S.: A security framework for e-Health service authentication and e-Health data transmission. In: Communications and Information Technology, 2009. ISCIT 2009. 9th International Symposium on, IEEE (2009) 1213–1218
19. Ryutov, T. and Zhou, L. and Neuman, C. and Leithead, T. and Seamons, K.E.: Adaptive trust negotiation and access control. In: Proceedings of the tenth ACM symposium on Access control models and technologies, ACM (2005) 139–146
20. Sandhu, R. and Ferraiolo, D. and Kuhn, R.: The NIST model for role-based access control: towards a unified standard. In: Symposium on Access Control Models and Technologies: Proceedings of the fifth ACM workshop on Role-based access control. Volume 26. (2000) 47–63
21. Berhe, S. and Demurjian, S. and Saripalle, R. and Agresta, T. and Liu, J. and Cusano, A. and Fequiere, A. and Gedarovich, J.: Secure, Obligated and Coordinated Collaboration in Health Care for the Patient-Centered Medical Home. In: AMIA Annual Symposium Proceedings. Volume 2010., American Medical Informatics Association (2010) 36
22. Caine, K. and Hanania, R.: Patients want granular privacy control over health information in electronic medical records. Journal of the American Medical Informatics Association **20**(1) (2013) 7–15
23. Bhatti, R. and Ghafoor, A. and Bertino, E. and Joshi, J.B.D.: X-GTRBAC: an XML-based policy specification framework and architecture for enterprise-wide access control. ACM Transactions on Information and System Security (TISSEC) **8**(2) (2005) 187–227
24. De la Rosa Algarín, A. and Demurjian, S. A., and Berhe, S. and Pavlich-Mariscal, J.: A Security Framework for XML Schemas and Documents for Health Care. In: International Workshop on Biomedical and Health Informatics. (2012) 782–789
25. Barrows, R.C. and Clayton, P.D.: Privacy, confidentiality, and electronic medical records. Journal of the American Medical Informatics Association **3**(2) (1996) 139–148
26. Silvey, SA and Schulte, J. and Smaltz, DH and Kamal, J. and others: Honest broker protocol streamlines research access to data while safeguarding patient privacy. In: AMIA... Annual Symposium proceedings/AMIA Symposium. AMIA Symposium. (2008) 1133
27. Buckovich, S.A. and Rippen, H.E. and Rozen, M.J.: Driving Toward Guiding Principles A Goal for Privacy, Confidentiality, and Security of Health Information. Journal of the American Medical Informatics Association **6**(2) (1999) 122–133